

DISCUSSION BEHAVIORS AND QUIZ PERFORMANCE AS A FUNCTION  
OF POINTS FOR DISCUSSION

An abstract of a thesis by  
Kathleen A. Colburn  
October 1976  
Drake University

The Problem: Social and point-token reinforcement were implemented to increase student question-asking and discussion behaviors.

Procedure: Three groups of university college students in small discussion classes were the participants in this study. Baseline data was collected on frequency of student discussion behaviors and on accuracy scores on both an immediate and delayed review quiz.

Findings: Results demonstrate that social and point-token reinforcement were highly effective for increasing the rate of discussion behaviors.

In addition, increased accuracy scores on the immediate review quiz following discussion of material to be contained in those quizzes were noted. No improvement on the delayed review quiz was observed.

Conclusion: Implications are that student interaction with academic material in the form of discussion can increase academic performance.

Recommendations: It would appear that efforts should be made to promote as much of this interaction as possible in the classroom.

DISCUSSION BEHAVIORS AND QUIZ PERFORMANCE AS A FUNCTION  
OF POINTS FOR DISCUSSION

---

A Thesis  
presented to  
The School of Graduate Studies  
Drake University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts

---

by  
Kathleen A. Colburn  
October, 1976

DISCUSSION BEHAVIORS AND QUIZ PERFORMANCE AS A FUNCTION  
OF POINTS FOR DISCUSSION

by

Kathleen A. Colburn

Approved by Committee:

John J. Williams

---

Chairman

William D. Klipec

---

Marjorie Prentice

Earle L. Canfield

---

Dean of the School of Graduate Studies

419327

## TABLE OF CONTENTS

	PAGE
INTRODUCTION AND REVIEW OF THE LITERATURE.....	1
METHOD.....	6
RESULTS.....	15
DISCUSSION.....	19
BIBLIOGRAPHY.....	26

# LIST OF FIGURES

FIGURE	PAGE
1. Rate of discussion behavior (freq./min.)	23
2. Frequency of discussion behaviors	24
3. Mean percentage accuracy score (Immediate review question)	25

## INTRODUCTION

The field of the applied analysis of behavior has been involved in various aspects of research on effective classroom management. In the elementary schools, investigators have suggested possible solutions for modifying disruptive behaviors (Bolstad & Johnson, 1972; Harris & Sherman, 1973; Madsen, Becker, & Thomas, 1969), increasing academic performance (Ayllon & Roberts, 1974; Ballard & Glynn, 1975; Felixbrod & O'Leary, 1973), and improving attending or on-task behaviors (Ferritor, Buckholdt, Hamblin, & Smith, 1972; Glynn & Thomas, 1974; Packard, 1970). The same effects have been achieved with a variety of different procedures such as individual contingencies for group consequences (Barrish, Saunders, & Wolf, 1969), token reinforcement (O'Leary, Becker, Evans, & Saudargas, 1969), and altering teacher attention to student behaviors (Broden, Bruce, Mitchell, Carter, & Hall, 1970).

Investigations have also been conducted on the high school level. One study (McAllister, Stachawiak, Baer, & Conderman, 1969) has demonstrated the usefulness of contingent teacher praise for increasing appropriate behaviors and disapproval for reducing problem behaviors such as talking out and turning around. The contingency that low grades (D's and F's) on quizzes in a high school French class would lead to individual tutoring after school resulted in marked improvement in those quizzes (Hall, Cristler, Cranston & Tucker, 1970). Other investigators have also demonstrated the usefulness of self-recording as a procedure for increasing the amount of time spent studying and decreasing the behavior of talking out in the classroom (Broden, Hall & Mitts, 1971).

In higher education, research has concentrated on procedures for producing better retention of material and improved test performance in the class.

Most of these investigations have centered around the use of personalized systems of instruction (PSI) in college classes (Keller, 1968). Some evaluative studies have assessed the relative contributions of the different components of PSI such as proctors, repeated testing, and study guides to the superior student performance and satisfaction in PSI courses while others have compared performance in PSI to traditional courses.

One study comparing the relative merits of a PSI course to a traditional lecture course produced data showing greater student satisfaction with the PSI course and superior performance on final examinations (McMichael & Corey, 1969). Other investigators found that a personalized course produced better student performance on written items in the unit quizzes. This finding is consistent with the results of another study which compared groups of students who received different percentages of proctored units from 0% to 100% (Farmer, Lachter, Blaustein, & Cole, 1972). The effect of proctoring was an acceleration of rate of progress through the course with better performance by the proctored students compared to the non-proctored ones.

Studies on the effect of frequent versus infrequent testing, as determined by length of study units, have shown mixed effects. Some studies have revealed that more frequent tests do not have any appreciable effects on exam performance but do appear to serve the function of regulating distribution of study time (Williams, 1975; Born, 1975). However, students in another study performed approximately 20% worse on a comprehensive course examination when they were required to master several unit quizzes at the same time (Semb, 1973). In a follow up study with long-assignment (infrequent tests) and short-assignment (frequent tests) groups, the long-assignment condition produced far more review quiz retakes than the short-assignment group, as well as

inferior test performance on first-attempt review quizzes and achievement tests (Semb, 1974). The achievement tests were a required short-essay test over each of the four parts of the course that did not contribute to the students' grade. A possible explanation for the difference in performance on these measures is that students in the long-assignment condition were not exposed to as many items from the test question pool because they did not have to take as many quizzes as the short-assignment group and this lack of interaction with the material could account for the different performance (Semb, 1974). Using the same line of reasoning, Bostow and O'Connor (1973) investigated the effects of allowing remediation of tests. Students allowed to remediate weekly quizzes received a higher average grade on a 100-item multiple choice comprehensive final than the no-remediation group. This finding is consistent with the argument that increased interaction with the material, in this case remediation, leads to superior performance.

Other studies have investigated the effects of interaction with the material through the use of study guides or interview procedures. One study demonstrated a 30% increase in performance on weekly quizzes simply by giving students study questions in advance (Semb, Hopkins, & Hursh, 1973). Sheppard and MacDermott (1970) used interview procedures as the method of teaching and found that students score significantly higher on objective and essay final examinations than students taught the same material with more conventional methods. The interview procedure required the student to describe fluently the material contained in the assigned section within a 10-minute period. The results from these two studies suggest that student interaction with the material, either verbal or written, will produce better retention of material as measured by examination or quiz performance.



Related findings have also been obtained in research outside the applied analysis of behavior. In a study on the long term effect of repetition, Bartz (1969) showed that subject restatement of presented stimulus items had its greatest facilitative effect in a delayed recall condition of 15 sec as opposed to an immediate recall condition. Participants who had the opportunity of adding repetitions of stimuli to the previous memory of such materials during the 15 sec demonstrated enhanced recall for these repeated lists in an additional final delayed recall test. In an investigation using a paired-associate task (noun-verb pair), Rohwer (1969) found that students who made up sentences using the words presented in the pair remembered them better than students who did not. Further investigation of this phenomena established that it was the behavior of making up the sentence that was important since others who read sentences made up by the experimenter did not recall as well as those who made up their own (Bobrow & Bower, 1969). This research from the verbal learning literature also suggests that verbal interaction with presented stimulus materials produces better recall.

In order to facilitate student verbal interaction with the material in the classroom, it is necessary to change the format from a traditional didactic lecture form to a discussion mode in which question asking and statements about the material are encouraged. However, simply allowing discussion to occur does not insure a high frequency of this behavior. A recent study by Evans (1970) provides a possible methodology for achieving this goal. He used social reinforcement to increase verbal behavior relevant to material in one of his classes. The investigation employed two classes to serve as two groups; the experimental group received social reinforcement for asking questions while questions from the control group were not reinforced. Results

showed that the experimental group asked significantly ( $\alpha = .05$ ) more questions over the same lectures than did the control. This study does not demonstrate that such an increase in question asking had an effect on academic performance or improved the students understanding of the material.

In an extension of Evans' study, Knapczyk and Livingston (1974) investigated the effects on students' performances as a result of training them to ask questions. They argued if students can be taught to request information from the teacher then this will provide feedback to the instructor regarding their level of understanding. Such feedback would enable the teacher to provide remedial instruction over difficult material and increase opportunities for students to learn. The more contact the student has with the material, the better he/she should perform on subsequent testing over that material. The participants in this study were two educable mentally retarded students who asked no questions in class prior to treatment. Training consisted of simply reminding the students to ask the teachers questions when either saw an unknown word. Their results demonstrated that as frequency of question asking increased, the percentage accuracy of reading performance increased also.

Both the verbal learning literature with the effects of repetition of presented stimuli and the applied analysis of behavior research with the use of study guides has shown that interaction with the material improves retention of that material. This research would suggest that increasing student interaction with classroom material would improve their retention of it. Increasing the frequency of question-asking and discussion statements in class would be a way of producing verbal interaction with the material. Point-token systems have been shown to have considerable strength in modifying a wide variety of behaviors (O'Leary et al., 1969; Kazdin & Bootzin, 1972). Such a procedure,

if applied contingently to discussion behaviors along with social reinforcement, would be expected to result in an increase in question asking and discussion (Evans, 1970).

If the frequency of discussion behaviors can be increased, then it is appropriate to ask whether this type of interaction with the material enhances performances measures. This study was designed to test both possibilities. Specifically, social reinforcement and a point contingency were established in an attempt to increase the frequency of discussion behaviors in relation to baseline levels. A second purpose was to evaluate whether such an increase would lead to improved performance on subsequent quiz questions over the class discussion both immediately following the class and one week later.

## METHOD

### Subjects

Students in three discussion sections of an introductory statistics course at Drake University were the subjects for this study. These students attended three one-hour lectures and the two hour discussion section per week. Sections were assigned as the first and second experimental groups and as the untreated control group according to the stability of the baseline data. There were eight discussion sections offered and students selected their own based on meeting time.

At the beginning of the study, the students were told they could change sections if they did not want to participate in the experiment. Twenty students were originally involved with two leaving, one because of a time conflict with

his job and the other stated he did not like the point contingency for discussion behaviors. Eighteen students participated throughout the entire study: eight men and one woman in the first experimental group, three men and one woman in the second experimental group, and four men and one woman in the control group.

### Setting and General Procedure

A student could earn a maximum of 200 points in the statistics course. The point requirements for the different letter grades was a minimum of 180 points for an A, 160 for a B, and 140 for a C. Half of the points were possible by taking four in-class tests and a comprehensive final, each worth 20 points. Forty points could be earned on four take-home tests worth 10 points each. The final 60 points could be earned at the rate of five points for each of the 12 discussion class meetings which provided the setting for the study.

The discussion leader met with each of the sections for two hours once a week. These meetings were designed to supplement and expand upon material covered in lectures by emphasizing the computational procedures of the statistics and discussing questions concerning the lecture material. During the first half hour the students took a quiz over the material presented during the previous week's discussion class and the reading assignment for the present week. The next one-and-one-half hours were spent discussing the material assigned for the week.

Three of the five points for each discussion class meeting were always available for performance on part of the weekly quizzes. The questions on those quizzes for these three points consisted of one computation and one

short answer essay over the previous week's computational exercise. The other two points available during each discussion class were contingent either on performance on four additional quiz questions or for discussion behaviors depending on the experimental condition as described below. When these points were awarded for quiz performance, the four questions required short essay answers based on the reading assignment for the present week.

At the first meeting of the semester, the students were informed that a study to identify some of the variables which might contribute to the effectiveness of a discussion class such as this one was being conducted. The instructor, who was the experimenter, read a list of instructions to the class which made the following points:

1. Two observers would be present during the discussion classes to record the verbal behavior of the instructor and students. Since the recording was to be scored by seat number to insure the students' anonymity, they were asked to sit in the same seat every week. Discussion behaviors were the only responses being recorded in class and the study involved no deception.

2. Later in the semester, specific discussion behaviors would be encouraged by giving points contingent on these behaviors. These points were those that were now available for performance on the reading quiz over the current week's assignment. The number of points involved were two per class. The other three points could always be earned for performance on quiz questions over the previous week's discussion material.

3. Each week the quiz would contain some questions which would not affect their grade.

4. Each week they would be given the number of points they had earned for the day when the class was finished.

5. If any student did not wish to participate, they could switch to one of the other five discussion sections not involved in the study.

6. Discussion was going to be encouraged. This would be facilitated by the instructor pausing 5 sec everytime one of the students made a statement or asked a question in order to provide the opportunity for other students to respond.

The instructor then asked if there were any questions about how the class was to be structured before discussion for the day began.

Except as noted below, the instructor attempted to maintain a zero rate of prompting and social reinforcement following a student verbal response. The observer, who also recorded instructor behavior (described below), provided an immediate hand signal as feedback whenever the instructor responded to student verbal behavior with a socially reinforcing comment or prompted discussion. This feedback procedure was implemented during baseline and reinforcement conditions for prompts but only during baseline for socially reinforcing comments.

### Measures

Discussion behaviors. Discussion behaviors were defined as: (a) any question relevant to the discussion topic for that class or previous classes or topics which would be covered in the future; and, (b) any statement relating to the discussion topic for that day or any other statistical concepts already covered or expected to be covered in the future. A relevant question was defined as any statement which pertained to the lecture material, the topic of

a relevant question or any statistical concept. A student's answer to any other student's question about the material as well as re-wording of material just presented was scored as a related statement. These responses were the principle dependent variable of the study. All other vocalizations such as questions about the mechanics of the course or material to be assigned for the quiz were considered irrelevant and were scored as described below.

Delayed review questions. The delayed review questions on each weekly quiz was a six-part question or, sometimes, two three-part questions covering the previous week's discussion material. It was added to the quiz along with the questions previously described but was not consequence with points nor was the identity of this non-contingent performance question revealed to the students until after the quiz. This measure began the second week of this study because there was no previous discussion material to quiz during the first week. Performance on these delayed review questions was the second dependent measure. They were designed to test the effects of discussion on retention and integration of the material after one week had elapsed.

Immediate review questions. The immediate review questions were a three-part short answer question administered at the end of each discussion class beginning at week seven. It covered the material just discussed in class that day which would also be contained in the delayed review question the next week. This review question was introduced to measure the effects of discussion on retention of material immediately following discussion because inspection of the data on the delayed review questions showed no consistent effect.

Instructor behavior. The observer scored two classes of instructor behavior, prompts and certain consequences of student verbal behavior.

Instructor prompts were defined as the instructor asking a question or requesting information from the class on the topic of discussion. The consequences scored were those that went beyond a simple answer or reply expected of an instructor to a student's question, especially those that might be considered socially reinforcing such as "good," "I'm glad you asked that," or "that's right." A simple reply was defined as the instructor providing factual information to a student's question such as "the mode is the score that occurs most often." Prompts were scored throughout the study. Consequences were scored only during the baseline phase of the study because social reinforcement was part of the reinforcement procedure.

#### Scoring and Reliability

Scoring. An independent observer recorded the number of discussion behaviors emitted by each student in each of the three discussion groups. Each discussion behavior was recorded on a pre-printed data sheet divided into 30 one-minute intervals for each student in the class. A frequency by interval recording system was used so that each instance of student verbal behavior was scored in the interval in which it occurred. The observers scored a "Q" for a relevant question, an "S" for a related statement, and "V" for all other vocalizations. Each time the instructor emitted a prompt a "?" was scored. Socially reinforcing consequences were recorded as a "C". These instructor behaviors were scored in the interval in which they occurred in the same manner as the student discussion behaviors. Generally the observers used three data sheets per discussion class on which they recorded both student and instructor behaviors.

The same observer corrected the answers and recorded the number of parts



each student answered correctly on the immediate review questions.

Reliability. Every week a separate observer scored student and instructor verbal behavior in at least two of the three discussion groups (except for week 10 when the observer was unavailable). This observer scored discussion behaviors for each individual in the discussion group and the number of prompts and reinforcing consequences emitted by the instructor in the same manner as the regular observer. The reliability observer also corrected both the immediate and delayed review questions. The instructor provided the observers with the answers to all the review questions.

Reliability was computed on the number of agreements between the two observers on the occurrence of each discussion behavior and on the number of agreements on the occurrence of each instructor prompt and/or reinforcing consequence. An agreement was defined as a discussion behavior or instructor behavior scored in the same interval by the two observers. Reliability on the immediate and delayed review questions was computed on the number of agreements between the two observers on scoring each part of each student's question correct or incorrect. The formula used for computing reliability on all the measures was the number of agreements over the number of agreements plus disagreements times 100.

#### Experimental Conditions

Baseline. In this condition, students discussion behaviors and instructor behavior were recorded. The instructor always paused for five sec after each student verbal response before continuing to speak. Delayed review questions were included on the quiz and, when the immediate review questions

were introduced in the study, these were presented at the end of the class. The two points, later used during the reinforcement condition, were contingent on correctly answering four questions over the reading assignment for the current week.

Reinforcement. In this condition, two sources of reinforcement were available. Immediate social reinforcement from the instructor was contingent on each discussion behavior and point-token reinforcement was contingent on emitting a specified number of discussion behaviors by the end of the class. Each time a student emitted a statement or question that was scored as relevant by the observer, a remote control response counter was operated by the observer to cue the instructor that she should reinforce the student. The instructor would then smile and praise the student with a statement like "good question" or "that's exactly right," while appearing genuinely spontaneous and pleased with the student's statement or question. After she had reinforced the student, the instructor would pause the five sec to allow other student responses to the question or statement just emitted.

The point token system consisted of two points contingent on emitting a criterion number of discussion behaviors during the whole class. This criterion for the first class of the reinforcement condition was defined for each student as a 50% increase over the median number of vocalizations during baseline. If the student met the criterion during the first session, it was raised again by another 50% each week until the criterion for each individual surpassed his/her highest data point in baseline. If the student did not meet the criterion, he/she was not awarded the two points and the criterion was reduced to either the median baseline level or, after the first class in this condition, to the last criterion met. Once the student met this requirement, it was

again adjusted upwards by the same procedure described previously. The two points used in this condition were those which had been contingent on the four questions over the reading assignment during baseline.

During this condition, the delayed review questions and the reading quiz questions were still included on the weekly quiz. The student were never told not to answer the reading quiz questions, they were just informed that the points they had received for answering these questions could now be earned for discussion behaviors in class. The immediate review questions were presented at the end of the class when they were introduced in the study.

During the week prior to the introduction of the reinforcement condition for each group, the students were informed of the changes that were to take place. They were read the following instructions:

Next week the format for the discussion group is going to change. Starting at the next meeting, the two points you earned for taking the reading quiz can now be earned for engaging in discussion either by asking a question, answering another student's question, or making a statement about the reading assignment or topic of discussion.

I have some definitions here of the types of discussion behaviors that will earn points. I'll give you a copy at the next meeting, but for now I'll just read them to you so you can get an idea of what will and will not earn points. (See discussion behavior section for definitions).

Also, at the next class meeting you will be given a small piece of paper on which will be written the number of questions or statements you are required to contribute to the discussion in order to earn the

two points for the day. So starting next class meeting you should be prepared to discuss the material.

At the beginning of every class during this condition, the students were handed a copy of the response definitions and a piece of paper with their criterion for the day written on it.

#### Sequence of Manipulations

A multiple baseline across two discussion groups with a reversal within one group was used to verify the independent variable (Baer, Wolf, & Risley, 1968; Knapczyk & Livingston, 1974). Once a stable baseline was achieved in the first group, the reinforcement procedure was implemented. When an increase in discussion behaviors had been clearly demonstrated, the baseline condition was reinstated. When the reversal had been demonstrated in Group 1 and a stable baseline evident for Group 2, the reinforcement procedure was introduced in Group 2 to increase the rate of discussion behaviors. Group 3, which served as the control group, remained on baseline throughout the entire study.

#### RESULTS

The reliability estimated for the measurement of instructor prompts and consequences ranged from 25% to 100% with agreement being 100% on all but the first week's observations. Median reliability on recording the discussion behaviors was 92% for Group 1 (range = 82% to 100%), 88% for Group 2 (range = 62% to 100% with two points below 80% early in the study), and 88% for Group 3 (range = 68% to 92% with one point below 80%). Median reliability results on

scoring the immediate review questions were 88% for Group 1 (range = 85% to 92%), 88% for Group 2 (range = 75% to 93%), and 90% for Group 3 (range = 84% to 92%). On the delayed review questions, the median reliability for Group 1 was 90% (range = 80% to 100%), 92% for Group 2 (range = 83% to 100%), and 96% for Group 3 (range = 90% to 100%).

The data on instructor behavior revealed that the instructor did not emit more than two prompts or two socially reinforcing consequences per session in any of the three discussion groups during their respective baseline conditions. During the reinforcement condition in Groups 1 and 2, the number of prompts emitted by the instructor was zero. Data was not kept on the number of socially reinforcing consequences during this condition because social reinforcement was part of the modification procedure.

One purpose of this study was to examine whether a point contingency would increase the rate of question asking and related statements that students emit in the classroom. Figure 1 depicts the rate of discussion behaviors for all three groups during the 12 discussion class meetings. Looking at the top panel, Group 1's rate varied between .05 and .38 discussion behaviors per minute during baseline. With the introduction of the point contingency, discussion ranged from .6 per minute to 1.07 per minute. The rate stabilized during weeks 7, 8, and 9 as all students emitted close to the minimum number of discussion behaviors required for reinforcement (with the one exception noted below). Removal of the points during the second baseline resulted in an immediate drop to .31 responses per minute at session 10 and ranged from .30 to .47. Group 2's rate of discussion behaviors (middle panel, Figure 1) ranged from .11 to .51 per minute during baseline, immediately rose to .71 per minute upon introduction of the point contingency at week 9, and ranged as high

as 1.17 at week 10. The data for Group 3 is presented in the bottom panel and shows no systematic change across the 12 weeks of the study. An analysis of variance utilizing the least squares method of regression (Winer, 1972), and the group as the unit of analysis, demonstrated that there was an overall effect across the entire study for differences in the frequency of discussion behaviors emitted by the three groups ( $F = 15.58$ ,  $df = 2/33$ ,  $p < .001$ ). Two a priori t-tests were conducted and revealed that the contrast between Group 1 during the reinforcement condition (weeks 5 through 9) and Group 3 during the same time period was significant ( $t = 4.20$ ,  $df = 33$ ,  $p < .001$ ), as was that between Group 2 and Group 3 during weeks 9 through 12 ( $t = 3.67$ ,  $df = 33$ ,  $p < .001$ ).

These within and between group findings were observed for all but one individual within the experimental groups. This individual in Group 1 stopped responding during weeks 7, 8, and 9 and consequently did not receive any points during these weeks. One other participant in Groups 1 and 2 failed to meet the criterion number of discussion behaviors on one occasion each. All other subjects in the study did reach their criteria each week. Each participant increased his/her frequency of discussion behaviors over baseline during the reinforcement condition although to varying degrees. That is, while the degree of change varied for each student, the direction of change was similar (see Figure 2).

Figure 2 presents the frequency of discussion behaviors for three representative students from Groups 1 and 2. Each of the three students' frequency of discussion behaviors varied according to the individual; however, when the point contingency was introduced the frequency of these behaviors increased above baseline levels. For student 1 and 2, discussion behaviors decreased in frequency when the point-token reinforcement condition was no

longer in effect (baseline 2).

Figure 3 shows the mean performance of each of the three groups on the immediate review questions. The top panel reveals that during the reinforcement condition, the mean percentage accuracy score of Group ranged from 82% to 93%. When the point contingency was removed, the mean percentage accuracy score declined sharply at weeks 10 and 11 to 56% and 50% respectively and then increased at the last class meeting. While Group 2 was in baseline (middle panel, weeks 7 and 8), the mean percentage accuracy score on the immediate review questions was 58% both weeks. With the introduction of the points during weeks 9 through 12, the mean score earned on these questions ranged from 50% to 92%. The bottom panel of Figure 3 shows the considerable variability and lack of any clear trend in the data for Group 3.

Unfortunately, time did not allow a reversal with either Group 1 or 2 on this measure. However, a statistical analysis was conducted to determine whether there was a reliable difference in the mean percentage accuracy scores between the groups over the entire study. The analysis revealed an overall difference in performance on the immediate review questions ( $F = 3.95$ ,  $df = 2/15$ ,  $p < .05$ ). Two a priori t-tests contrasting the performance of each experimental group with the control group revealed a significant difference between Group 1 and Group 3 during weeks 7, 8, and 9 ( $t = 2.39$ ,  $df = 15$ ,  $p < .025$ ) and a difference approaching significance between Group 2 and Group 3 during weeks 9 through 12 ( $t = 1.488$ ,  $df = 15$ ,  $p < .10$ ).

Inspection of the delayed review questions data revealed no clear effect of the reinforcement condition on scores of the two experimental groups. However, since there was a suggestion of an effect with Group 2, a statistical analysis was conducted but did not allow rejection of the null hypothesis.

## DISCUSSION

The results suggested that the point contingency with social reinforcement effectively increased the rate of discussion behaviors. Since reliability on all dependent measures was very high, it is assumed that these results are dependable. The reinforcement package of the present study seemed well suited to the educational environment because it could be easily implemented in the traditional classroom. It taught an educationally relevant response which provided feedback to the instructor on areas where the students needed additional information. In a classroom situation it is important that the student be provided with the opportunity and incentive to clarify as much information about the subject matter as necessary. The functional effect of this reinforcement package was to assure that students would take advantage of this opportunity.

Except for the one subject mentioned in the results section who failed to respond to the contingencies, the group curves were representative of individual performance on all three measures. These results provided evidence that the reinforcement had a powerful effect on nearly every students' rate of discussion.

The significant difference found on the accuracy scores for the immediate review questions between Group 1 during reinforcement for discussion behaviors and Group 3 suggests that increasing discussion, thereby increasing interaction with the subject matter, was responsible for these higher scores. However, another explanation is possible. It could be argued that the higher rate of reinforcement received by Group 1 was directly responsible for the increased accuracy scores rather than the interaction with the material (discussion) being the effective variable. Other research has demonstrated though that, compared to contingent reinforcement, non-contingent reinforcement does not result in



increased frequency of a targeted behavior and, sometimes, results in a decrease in responding (Goetz, Holmberg, & LeBlanc, 1975; Smeets & Striefel, 1975).

Since reinforcement in this study was contingent on an increase in discussion behaviors and was non-contingent with respect to accuracy on the immediate review questions, it is most likely that the increase in discussion behaviors in Group 1 was responsible for the higher accuracy on these questions.

Failure to achieve a statistically significant difference between Groups 2 and 3 on the immediate review questions is perhaps best explained by the very small number of subjects in Group 2 ( $n = 4$ ). It is clear that, except for week 10, there was a substantial increase in the students accuracy scores when reinforcement for discussion was in effect.

Overall, it appears that increasing discussion behaviors with the reinforcement condition of this study does lead to retention of the material as measured by accuracy of performance on quizzes administered immediately after the discussion has occurred. This effect is probably due to remedial and/or more extensive information provided by the instructor as difficult areas for students were observed. The same information was sometimes presented in three or four different modes or examples in order to clarify a point. The effect of these changes in the behavior of both the instructor and students was increased accuracy on the immediate review questions.

The delayed review question results show that neither of the experimental groups demonstrated superior performance on these questions over the control group. Forgetting of the material during the week appeared to be similar for all three groups despite the contingencies encountered in the classroom the previous week. It is difficult to explain the lack of results but a number of issues made it arduous to produce clear changes on this measure. Brigham,

Finfrock, Breunig, & Bushell (1972) have argued that it is difficult to consistently affect academic dependent variables because variation in the complexity of most academic materials confounds analysis of contingencies. Changes in the required response unit each week probably had more effect on the students' performance than the reinforcement received the previous week. Another issue which faces all behavior analysts conducting research in the area of education is that academic behaviors are so complex that it is nearly impossible to control all the variables that effect them (Ulrich, 1975).

Research from the PSI and verbal learning literature suggests that written and/ or verbal interaction with the material improves retention of it. The present study demonstrated some findings consistent with this body of literature. Increased interaction with the material in the form of discussion improved accuracy scores on an immediate test of retention but not on a similar test one week later. The reinforcement condition in this study was not powerful enough to indirectly effect a performance variable measured a week after the reinforcement contingencies were encountered. One reason is the difficulty of the material varied over the semester (i.e. from descriptive to inferential statistics to chi-square) and, as suggested by Brigham et al, these variations in the material probably overrode the reinforcement in the classroom. However, as interaction with the subject matter did improve accuracy scores for the immediate review questions, more interaction might be beneficial for improving performance on the delayed review questions. Future research should be devoted to investigating the effects of innovative methods of increasing student interaction with the subject matter in the classroom on a delayed performance measure.

Discussion behaviors are an educationally relevant response that should be

learned by all students as a prerequisite to a move from the didactic to the discussion method of learning. This method of learning appears to improve accuracy on some academic measures and provides a setting for exchange of information between the student and teacher. For these reasons it would seem important to have these behaviors in every student's academic repertoire. The reinforcement procedure used in this study is one way of insuring that such behaviors become a part of that repertoire.

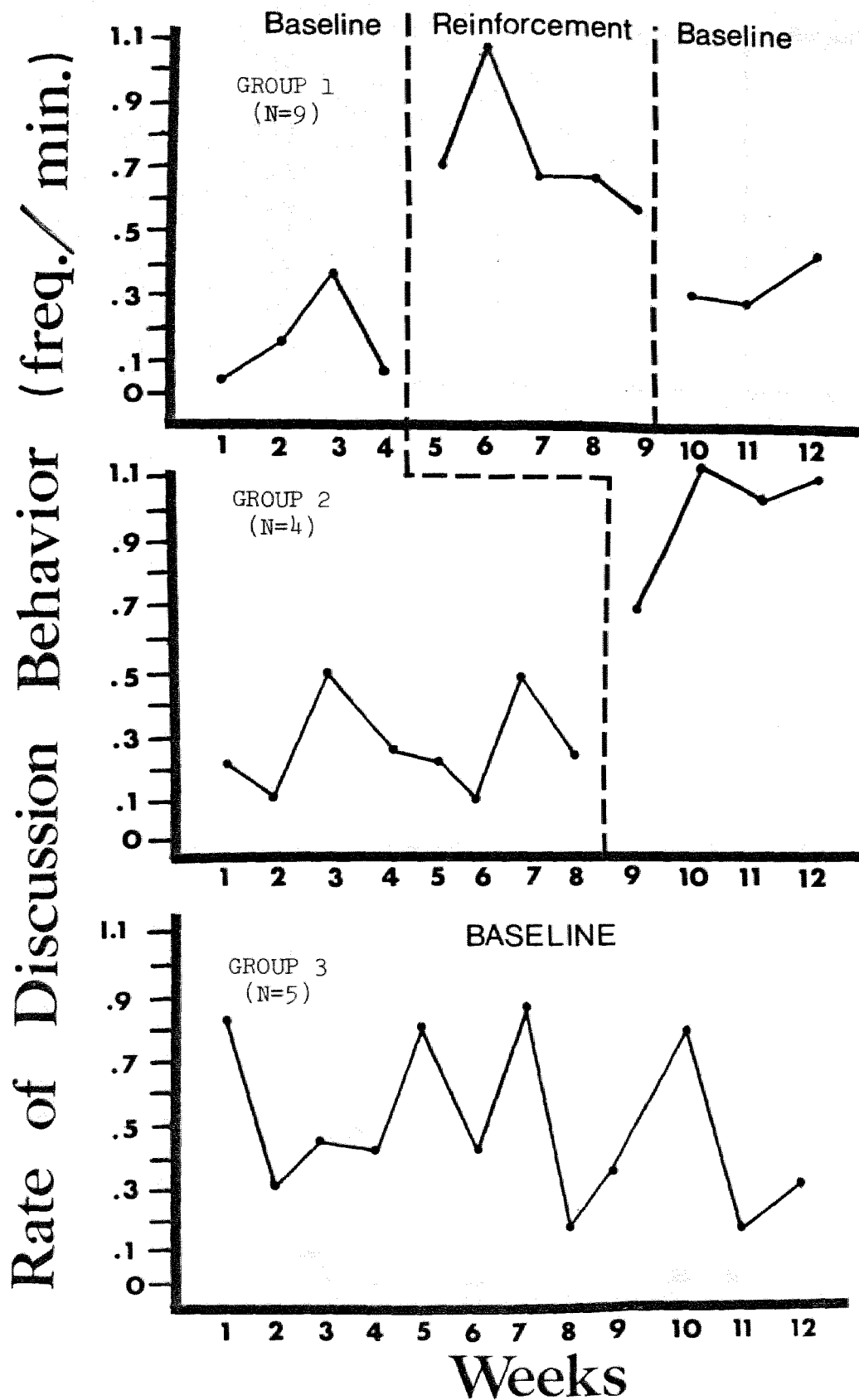


Fig. 1. Rate of discussion behaviors (responses/minute) during baseline and reinforcement conditions for Groups 1, 2, and 3.

# Frequency of Discussion Behaviors

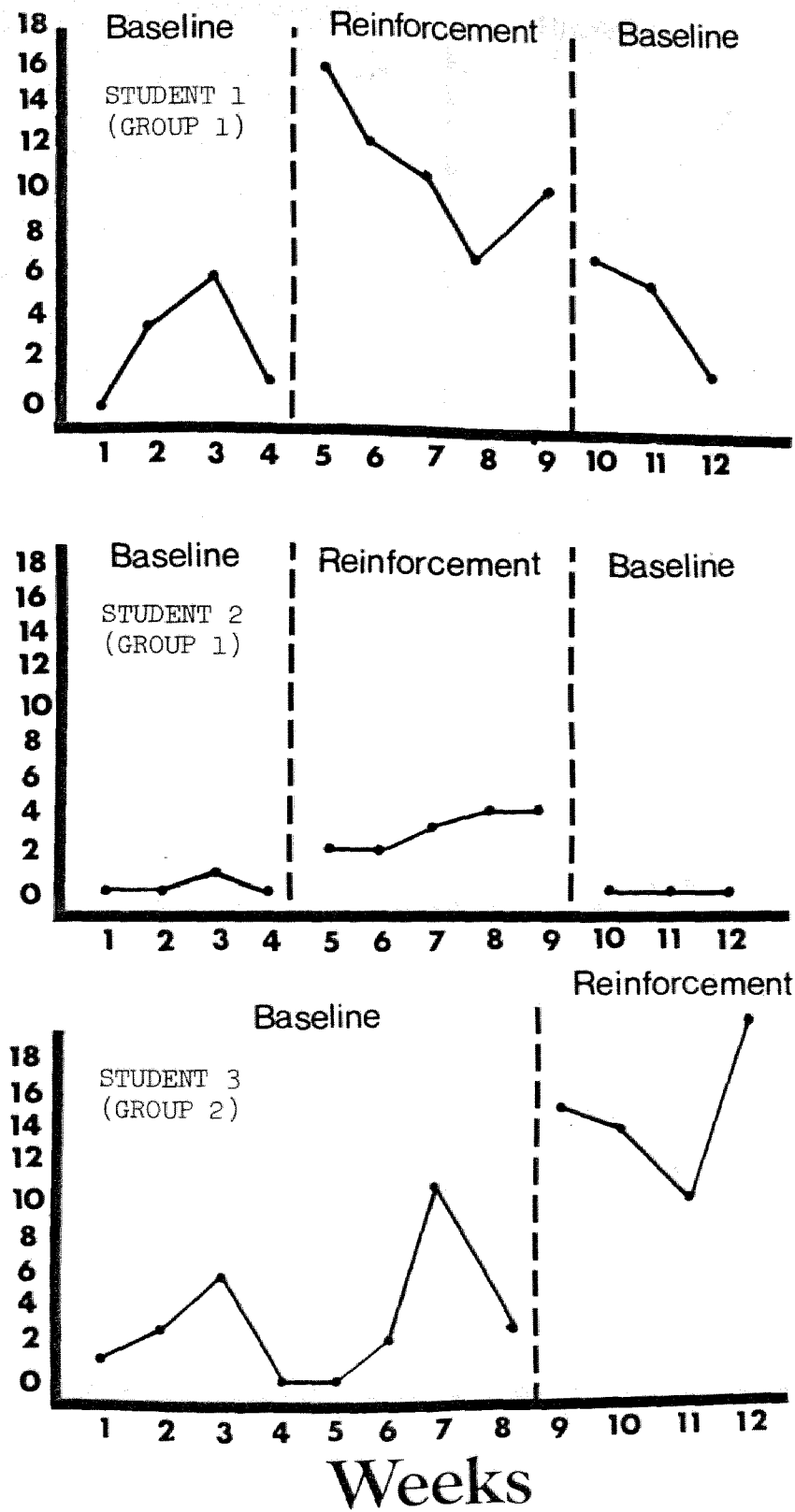


Fig. 2. Frequency of discussion behaviors emitted by three students during baseline and reinforcement conditions.

# Mean Percentage Accuracy Score (IMMEDIATE REVIEW QUESTIONS)

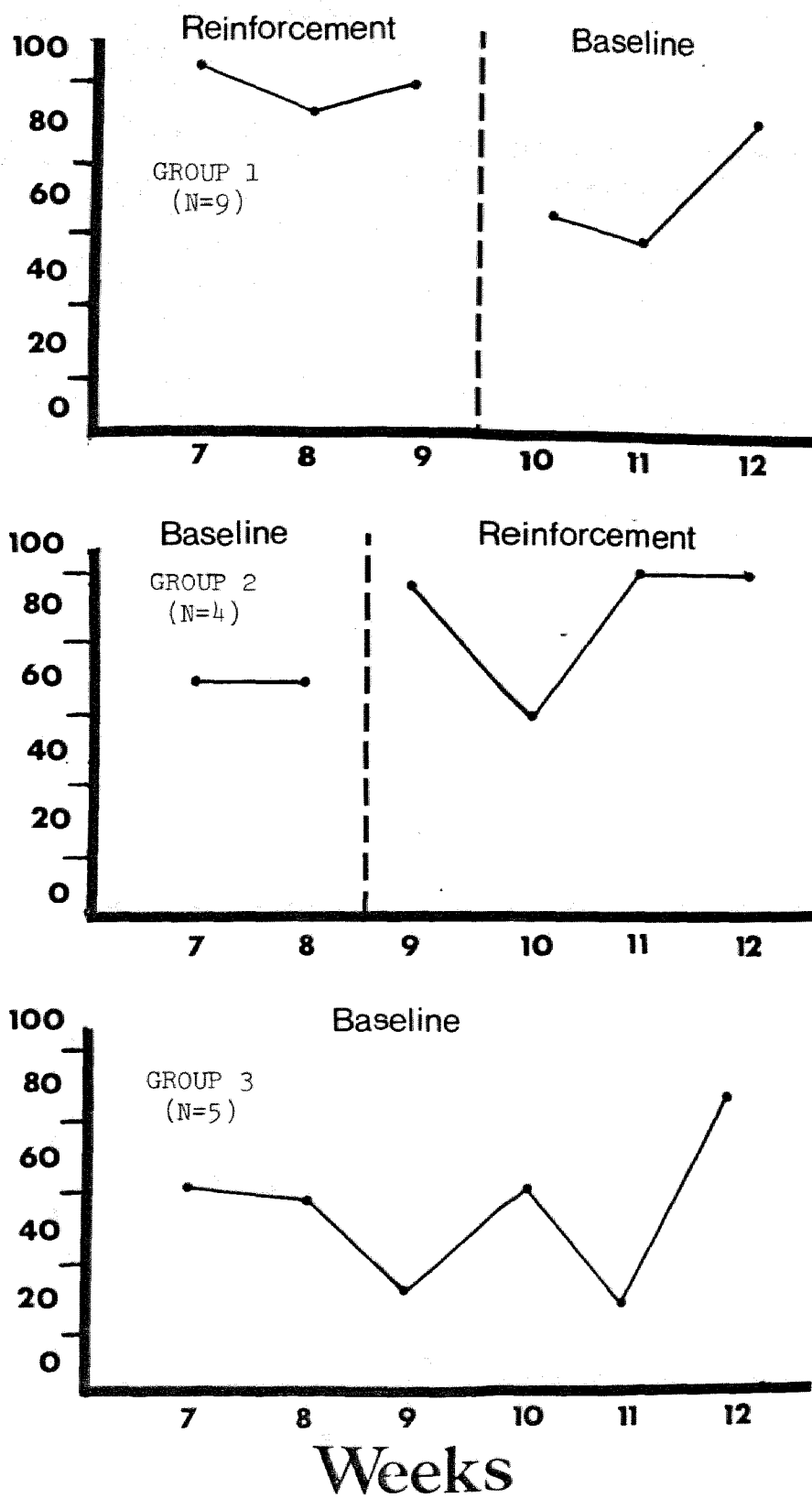


Fig. 3. Mean percentage accuracy score on immediate review questions as a function of giving points for discussing material contained in these questions.

## REFERENCES

- Ayllon, T., & Roberts, M. D. Eliminating discipline problems by strengthening academic performance. Journal of Applied Behavior Analysis, 1974, 7, 71-76.
- Baer, D. M., Wolf, M. M., & Risley, T. R. Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1968, 1, 91-97.
- Ballard, K. D. & Glynn, T. Behavioral self-management in story writing with elementary school children. Journal of Applied Behavior Analysis, 1975, 8, 387-398.
- Barrish, H. H., Saunders, M., & Wolf, M. M. Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. Journal of Applied Behavior Analysis, 1969, 2, 119-124.
- Bartz, W. H. Repetition and the memory stores. Journal of Experimental Psychology, 1969, 80, 33-38.
- Bobrow, S. A., & Bower, G. H. Comprehension and sentence recall. Journal of Experimental Psychology, 1969, 80, 455-461.
- Bolstad, O. D., & Johnson, S. M. Self-regulation in the modification of disruptive classroom behavior. Journal of Applied Behavior Analysis, 1972, 5, 443-454.
- Born, D. G. Exam performance and study behavior as a function of study unit size. In J. M. Johnston (Ed.) Behavior research and technology in higher education. Springfield, Ill.: Thomas, 1975.
- Born, D. G., Gledhill, S. M., & Davis, M. L. Examination performance in lecture discussion and personalized instruction courses. Journal of Applied Behavior Analysis, 1972, 5, 33-43.
- Bostow, D. E., & O'Connor, R. J. A comparison of two college classroom testing procedures: required remediation versus no remediation. Journal of Applied Behavior Analysis, 1973, 6, 599-607.
- Brigham, T. A., Finfrock, S. R., Breunig, M. K., & Bushell, D. The use of programmed materials in the analysis of academic contingencies. Journal of Applied Behavior Analysis, 1972, 5, 177-182.
- Broden, M., Bruce, C., Mitchell, M. A., Carter, V., & Hall, R. V. Effects of teacher attention on attending behavior of two boys at adjacent desks. Journal of Applied Behavior Analysis, 1970, 3, 199-203.
- Broden, M., Hall, R. V., & Mitts, B. The effect of self-recording on the classroom behavior of two eighth grade students. Journal of Applied Behavior Analysis, 1971, 4, 191-199.
- Evans, D. R. Social reinforcement of question asking behavior. Western Psychologist, 1970, 2, 80-83.

- Farmer, J., Lachter, A. D., Blaustein, J. J., & Cole, B. K. The role of proctoring in personalized instruction. Journal of Applied Behavior Analysis, 1972, 5, 401-404.
- Felixbrod, J. J., & O'Leary, K. D. Effects of reinforcement on children's academic behavior as a function of self-determined and externally imposed contingencies. Journal of Applied Behavior Analysis, 1973, 6, 241-250.
- Ferritor, D. E., Buckholdt, D., Hamblin, R. L., & Smith, L. The noneffects of contingent reinforcement for attending behavior on work accomplished. Journal of Applied Behavior Analysis, 1972, 5, 7-17.
- Glynn, E. L., & Thomas, J. D. Effect of cueing on self-control of classroom behavior. Journal of Applied Behavior Analysis, 1974, 7, 299-306.
- Goetz, E. M., Holmberg, M. C., & LeBlanc, J. M. Differential reinforcement of other behavior and noncontingent reinforcement as control procedures during the modification of a preschooler's compliance. Journal of Applied Behavior Analysis, 1975, 8, 77-82.
- Hall, R. V., Cristler, C., Cranston, S. S., & Tucker, B. Teachers and parents as researchers using multiple baseline designs. Journal of Applied Behavior Analysis, 1970, 3, 247-255.
- Harris, V. W., & Sherman, J. A. Effects of peer tutoring and consequences on the math performance of elementary classroom students. Journal of Applied Behavior Analysis, 1973, 6, 587-597.
- Kazdin, A. E., & Bootzin, R. R. The token economy: an evaluative review. Journal of Applied Behavior Analysis, 1972, 5, 343-372.
- Keller, F. A. Good-bye teacher....Journal of Applied Behavior Analysis, 1968, 1, 79-89.
- Knapczyk, D. R., & Livingston, G. The effects of prompting question-asking upon on-task behavior and reading comprehension. Journal of Applied Behavior Analysis, 1974, 7, 115-121.
- Madsen, C. J. Jr., Becker, W. C., & Thomas, D. R. Rules, praise, and ignoring: elements of elementary classroom control. Journal of Applied Behavior Analysis, 1969, 2, 277-285.
- McAllister, L. W., Stachowiak, J. G., Baer, D. M., & Conderman, L. The application of operant conditioning techniques in a secondary school classroom. Journal of Applied Behavior Analysis, 1969, 2, 277-285.
- McMichael, J. S., & Corey, J. R. Contingency management in an introductory psychology course produces better learning. Journal of Applied Behavior Analysis, 1969, 2, 79-83.
- O'Leary, K. D., Becker, W. C., Evans, M. B., & Saudargas, R. A. A token reinforcement program in a public school: a replication and systematic analysis. Journal of Applied Behavior Analysis, 1969, 2, 3-13.



- Packard, R. G. The control of "classroom attention:" a group contingency for complex behavior. Journal of Applied Behavior Analysis, 1970, 3, 13-28.
- Rohwer, W. D., Jr. Constraint, syntax, and meaning in paired-associate learning. Journal of Verbal Learning and Verbal Behavior, 1969, 5, 541-547.
- Semb, G. Human development lecture notes and study guide. Lawrence, Kansas: Department of Human Development, University of Kansas, 1973.
- Semb, G. The effects of mastery criteria and assignment length on college student test performance. Journal of Applied Behavior Analysis, 1974, 7, 61-69.
- Semb, G., Hopkins, B. L., & Hursh, D. E. The effects of study questions and grades on students test performance in a college course. Journal of Applied Behavior Analysis, 1973, 6, 631-643.
- Sheppard, W. C., & MacDermott, H. G. Design and evaluation of a programmed course in introductory psychology. Journal of Applied Behavior Analysis, 1970, 3, 5-11.
- Smeets, P. M., & Striefel, S. The effects of different reinforcement conditions on the test performance of multi-handicapped deaf children. Journal of Applied Behavior Analysis, 1975, 8, 83-89.
- Ulrich, R. Toward Experimental Living, Phase II: Have You Ever Heard of a Man Named Frazier, Sir? In E. Ramp & G. Semb (Eds.) Behavior analysis: Areas of research and application. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1975.
- Williams, R. L. The combined and differential effects of two types of preparatory quizzes on college student behaviors. In J. M. Johnston (Ed.) Behavior research and technology in higher education. Springfield, Ill.: Thomas, 1975.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Company, 1972.